

An Integrated Analysis of LOTS–HOTS Distribution and Instructional Alignment in Indonesian Language Final Examination at SMP Negeri 2 Dayun

Agnessya Juliana¹, Alber¹

¹ Universitas Islam Riau, Pekanbaru, Indonesia
agnessyajuliana@student.uir.ac.id, alberuir@edu.uir.ac.id
corresponding author: agnessyajuliana@student.uir.ac.id

ABSTRACT

This study aims to analyze the distribution of Lower Order Thinking Skills (LOTS) and Higher Order Thinking Skills (HOTS), as well as the alignment between test items and learning objectives in the odd-semester Final Examination (UAS) in the Indonesian Language subject at Grade VII at SMP Negeri 2 Dayun. This study employed a descriptive qualitative approach using content analysis of 50 examination items, answer keys, learning objectives, and interview data from the Indonesian Language teacher. The analysis was conducted based on the revised Bloom's Taxonomy by Anderson, Krathwohl, and Webb's alignment principles. The findings reveal that the examination is predominantly oriented toward LOTS, with 47 items (94%) categorized as LOTS and only 3 items (6%) classified as HOTS, mainly at the cognitive level C3 (applying). In addition, although 64% of the items were aligned with the learning objectives, 36% showed inconsistencies in cognitive demand, indicating that several test items did not adequately reflect the intended learning outcomes. This imbalance demonstrates that the assessment has not fully supported the development of students' higher-order thinking skills, particularly critical, analytical, and creative thinking. The study highlights the need for more balanced and cognitively challenging assessment practices through the integration of HOTS-oriented items and stronger alignment between assessment instruments and instructional objectives. Scientifically, this study provides an evaluative framework for examining both cognitive level distribution and item alignment, which can serve as a reference for improving assessment quality in accordance with the principles of the Merdeka Curriculum.

Keywords: LOTS, HOTS, item alignment, learning objectives, assessment, Indonesian language

Submitted	Accepted	Published
22 April 2026	07 May 2026	30 May 2026

Citation	:	Juliana, A., & Alber, A. (2026). An Integrated Analysis of LOTS–HOTS Distribution and Instructional Alignment in Indonesian Language Final Examination at SMP Negeri 2 Dayun. <i>Jurnal PAJAR (Pendidikan dan Pengajaran)</i> , 10(3), 335-346. DOI: http://dx.doi.org/10.33578/pjr.v10i3.1509 .
-----------------	---	--

INTRODUCTION

The Merdeka Curriculum positions assessment as an essential component in determining the success of learning. Assessment functions not only as a tool to measure students' mastery of content but also as a means to continuously monitor the development of their thinking processes. Through assessment, teachers can identify students' abilities to understand, apply, analyze, and process information critically and reflectively. Therefore, assessment instruments must be designed to accurately represent variations in students' cognitive abilities and to support the development of higher-order thinking skills in accordance with curriculum demands.

One form of assessment used to measure learning achievement is the Final Semester Examination (UAS). The Ministry of Education, Culture, Research, and Technology (2022) states that the UAS is a summative assessment conducted at the end of the learning process to comprehensively evaluate students' learning outcomes. Consequently, the quality of UAS test items greatly influences the validity of the information obtained regarding students' competencies and learning achievement over one semester. In this context, assessment instruments should not merely focus on factual recall but must also encourage analytical, evaluative, and creative thinking processes.

Based on the perspective of educational assessment, assessment is understood as a systematic process to determine the level of achievement of learning objectives through the measurement of learning outcomes. Sukardi (2008) emphasizes that assessment instruments must be aligned with learning objectives in order to validly measure the targeted competencies. Similarly, Webb (1997) explains that alignment between learning

objectives and assessment is essential to ensure consistency between intended competencies and measured outcomes. This alignment includes two important dimensions, namely content alignment and cognitive level alignment. Test items are considered valid when the material and cognitive demands correspond to the competencies formulated in the learning objectives.

In Indonesian language learning, multiple-choice questions are widely used because they provide objectivity and consistency in scoring (Margareta et al., 2024). Nevertheless, the quality of assessment is not determined by the item format itself, but by the cognitive processes required to answer the questions (Yulistio & Nugroho, 2024). Therefore, the analysis of cognitive levels in test items becomes crucial to identify whether assessment instruments have facilitated the measurement of both lower-order and higher-order thinking skills.

According to Anderson and Krathwohl (2015), cognitive abilities are classified into six levels: remembering, understanding, applying, analyzing, evaluating, and creating. Nugroho (2018) further categorizes these levels into Lower Order Thinking Skills (LOTS), which include C1–C3, and Higher Order Thinking Skills (HOTS), which include C4–C6. HOTS-oriented assessments are essential in the Merdeka Curriculum because they encourage students to think critically, solve problems, and construct knowledge independently. Therefore, assessment instruments should contain a proportional distribution of cognitive levels to comprehensively measure students' cognitive competencies (Abraham et al., 2021).

This study was conducted at SMP Negeri 2 Dayun, a school that has implemented the Merdeka Curriculum in its learning process. In Grade VII Indonesian language learning, students are expected not only to understand various types of texts but also to analyze text structures, evaluate information, and communicate ideas critically in oral and written forms. These competencies indicate that the learning objectives require cognitive processes extending beyond memorization and understanding, thereby demanding assessment instruments capable of measuring higher-order thinking skills.

Preliminary interviews with the Grade VII Indonesian language teacher reveal that the preparation of UAS test items was based on learning objectives, levels of difficulty, and reference materials from textbooks and digital resources. However, preliminary observations show that numerous test items still focused on lower cognitive processes such as identifying and recalling information. Several items merely required students to recognize text characteristics at the remembering level (C1), while the learning objectives targeted applying or analyzing skills (C3–C4). In addition, mismatches were identified between the cognitive demands of the learning objectives and the cognitive levels measured in the assessment items. This condition indicates that the implementation of cognitive alignment principles in assessment construction has not been fully optimized.

Previous studies consistently demonstrate that school assessments are still dominated by LOTS-oriented questions. Benny Antari et al. (2021) found that 92% of Mathematics test items were concentrated at the C3 level. Wahyuni et al. (2022) also reported a dominance of C3 and C4 levels with very limited proportions of C5 and C6. In Indonesian language learning, Nurrohmah and Mukhlis (2023) identified a dominance of C1-level questions without any HOTS items. Similar findings were reported by Benu et al. (2024), who found that 60% of the items were categorized at the C2 level, while Karo-Karo et al. (2025) showed that 97.14% of the items remained within the C1–C3 categories. Although these studies provide essential information regarding cognitive-level distribution, most of them focus only on classifying LOTS and HOTS without examining whether the cognitive demands of assessment items are aligned with the intended learning objectives. As a result, empirical studies integrating cognitive-level analysis with alignment analysis in the context of the Merdeka Curriculum remain limited.

Based on this gap, the present study offers a more comprehensive approach by simultaneously analyzing the distribution of LOTS and HOTS and examining the alignment between test items and learning objectives in terms of both content and cognitive level. Unlike previous studies that mainly focused on the proportion of cognitive categories, this study investigates how inconsistencies between learning objectives and assessment demands may affect the validity of learning evaluation. Thus, the novelty of this study lies in

the development of an integrated evaluative framework that combines cognitive taxonomy analysis with alignment analysis to identify not only the dominance of cognitive levels but also the extent to which assessment practices reflect curriculum expectations. This approach provides a more in-depth empirical contribution to the field of educational evaluation, particularly in strengthening assessment validity and improving the quality of HOTS-oriented assessment design within the implementation of the Merdeka Curriculum.

METHOD

This study employs a qualitative descriptive approach aimed at systematically examining the cognitive process levels and the alignment between test items and learning objectives in the odd-semester Final Examination (UAS) of Grade VII Indonesian Language at SMP Negeri 2 Dayun. This approach was selected because the study focuses on describing and interpreting assessment phenomena based on authentic educational documents rather than testing hypotheses or manipulating variables. In addition, the qualitative descriptive approach enables an in-depth interpretation of the cognitive demands reflected in the test items and their consistency with the intended learning objectives within the implementation of the Merdeka Curriculum.

This study uses the content analysis method because it allows documents to be analyzed systematically, objectively, and contextually in order to identify cognitive patterns and instructional alignment. Previous studies on LOTS and HOTS assessments have generally focused only on the percentage distribution of cognitive levels without employing explicit coding procedures or alignment analysis frameworks. Methodologically, such studies rarely explain the unit of analysis, coding stages, or inter-rater reliability procedures in detail. Therefore, this study adopts a more systematic analytical procedure by integrating cognitive classification analysis with alignment analysis using explicit coding indicators based on the Revised Bloom's Taxonomy and Webb's alignment principles.

The data sources consist of official educational documents, namely the odd-semester Final Examination questions for Grade VII Indonesian Language, answer keys, and learning objective documents developed by the teacher at SMP Negeri 2 Dayun. These documents were selected because they represent authentic evidence of assessment practices and instructional planning implemented within the classroom context. Supporting data were also obtained through semi-structured interviews with the Indonesian Language teacher to gain information regarding the process of test construction, consideration of cognitive levels, and the teacher's understanding of assessment alignment.

The unit of analysis in this study is each individual test item contained in the examination document. A total of 50 multiple-choice items were analyzed, including question items, answer options, and their corresponding learning objectives. Each item was examined based on the operational verbs, contextual demands, thinking processes, and compatibility with the intended learning objectives. The analysis focused on identifying both the cognitive level and the alignment between assessment indicators and instructional objectives.

To ensure analytical consistency, this study employed a cognitive classification rubric adapted from the Revised Bloom's Taxonomy by Anderson and Krathwohl (2015). The rubric was used as the primary instrument for coding cognitive levels in the test items. The classification indicators are presented in the following table.

Table 1. Cognitive Indicator

Cognitive Levels	Indicators of Cognitive Process	Categories
C1 Remembering	Identifying, recalling, and mentioning facts or information	LOTS
C2 Understanding	Explaining, interpreting, and summarizing meanings or concepts	LOTS
C3 Applying	Using concepts or procedures in specific situations	LOTS

C4 Analyzing	Differentiating, organizing, and connecting information critically	HOTS
C5 Evaluating	Assessing, criticizing, and justifying based on criteria	HOTS
C6 Creating	Producing, designing, and constructing new ideas or texts	HOTS

The coding process was conducted in several stages. First, the researcher identified operational verbs and cognitive demands in each test item. Second, each item was coded into cognitive categories C1–C6 based on the established rubric. Third, the coded items were grouped into LOTS and HOTS categories. Fourth, alignment analysis was carried out by comparing the cognitive demands and content of each item with the corresponding learning objectives. The alignment analysis referred to Webb’s principles, which emphasize consistency between learning objectives and assessment in terms of both content relevance and cognitive complexity. The results of coding and alignment analysis were then tabulated systematically to identify cognitive distribution patterns and alignment tendencies.

Data analysis followed the interactive analysis model consisting of data reduction, data display, and conclusion drawing. During data reduction, irrelevant information was eliminated while cognitive indicators and alignment characteristics were identified. The data were subsequently displayed in analytical tables containing item numbers, cognitive levels, LOTS/HOTS categories, and alignment status. The final stage involved interpreting the findings to explain patterns of cognitive distribution and the extent of alignment between assessment instruments and learning objectives.

To strengthen the validity and trustworthiness of the findings, this study employed expert validation and inter-rater reliability procedures. The coding results and alignment classifications were reviewed independently by an expert in educational evaluation and assessment. The expert examined the appropriateness of cognitive categorization, the interpretation of operational verbs, and the consistency of alignment analysis. After following the independent review process, discrepancies between the researcher and the expert were discussed until agreement was reached. This procedure was conducted to enhance confirmability, consistency, and analytical reliability, ensuring that the findings are scientifically valid and accountable.

RESULTS AND DISCUSSION

Distribution of LOTS and HOTS in Test Items

The analysis of the Grade VII Indonesian Language Final Semester Examination (UAS) items at SMP Negeri 2 Dayun was conducted using the Revised Bloom’s Taxonomy proposed by Anderson and Krathwohl, which classifies cognitive processes into six levels: remembering (C1), understanding (C2), applying (C3), analyzing (C4), evaluating (C5), and creating (C6). These cognitive levels were subsequently grouped into Lower-Order Thinking Skills (LOTS) and Higher-Order Thinking Skills (HOTS). The analysis aimed to identify the extent to which the assessment instrument reflects the cognitive demands emphasized in the Merdeka Curriculum.

Table 2. Percentage of Cognitive Level

Cognitive Levels	Categories	Number of Items	Percentages
C1 Remembering	LOTS	11	22%
C2 Understanding	LOTS	15	30%
C3 Applying	LOTS	21	42%
C4 Analyzing	HOTS	3	6%
C5 Evaluating	HOTS	0	0%
C6 Creating	HOTS	0	0%
Total		50	100%

As displayed above, the majority of the test items are categorized as LOTS, with 47 items (94%) distributed across levels C1, C2, and C3. Meanwhile, only 3 items (6%) are classified as HOTS at the C4 level, while no items are identified at the evaluating (C5) and creating (C6) levels. These findings indicate that the examination instrument predominantly measures basic cognitive processes related to remembering, understanding, and procedural application.

The dominance of LOTS-oriented questions demonstrates that the assessment still prioritizes factual mastery and conceptual comprehension rather than encouraging students to engage in critical and reflective thinking. In the context of the Merdeka Curriculum, this imbalance is significant because the curriculum emphasizes competency-based learning and the development of higher-order thinking skills. Consequently, assessments that are overly concentrated at lower cognitive levels may not adequately support students' analytical and creative capacities.

This finding is consistent with previous studies showing that classroom assessments in Indonesia remain largely focused on LOTS. Nurrohmah and Mukhlis (2023) found that Indonesian language assessments were dominated by C1-level questions and lacked HOTS representation. Similarly, Karo-Karo et al. (2025) reported that 97.14% of assessment items were concentrated at C1–C3 levels. International research also reveals similar tendencies. Aydin and Birgili (2023) argue that school assessments frequently emphasize procedural mastery rather than complex reasoning, resulting in limited opportunities for students to develop higher-order thinking skills. Likewise, Zhou et al. (2023) emphasize that balanced cognitive distribution is necessary to promote meaningful learning and critical engagement.

Representation of LOTS in Test Items

The findings indicate that LOTS items dominate the assessment instrument, particularly at the C3 (applying) level. Although application tasks involve a higher cognitive demand than remembering and understanding, the questions remain procedural and structured, limiting students' opportunities to critically evaluate information or generate original ideas.

One representative example of the C1 (remembering) level is Item 21:

“Which of the following represents the structure of a Pantun?”

This item requires students only to recognize the characteristics of a Pantun that have been previously learned. Students are not expected to interpret, compare, or analyze the structure critically; instead, they simply retrieve factual information from memory. The cognitive demand is therefore limited to recognizing and recalling information stored in long-term memory. According to Anderson and Krathwohl, remembering tasks focus solely on retrieving previously acquired knowledge without requiring transformation or deeper processing.

The prevalence of such questions indicates that the assessment still prioritizes memorization-oriented learning. While remembering-level items are essential for measuring foundational knowledge, excessive reliance on these tasks may reduce opportunities for students to engage in more complex thinking processes. This condition reflects a traditional assessment orientation in which success is measured primarily through factual recall.

At the C2 (understanding) level, the questions generally require students to interpret textual meaning and identify implicit information. For example, item 1 states:

“The main purpose of the author in the text ‘The Instagrammable Pantan Terong’ is...”

To answer this question, students must comprehend the text and identify the communicative intention of the author. The cognitive process includes interpreting and inferring meaning from contextual information. Unlike C1 tasks, this item requires students to construct meaning from the text rather than merely recalling facts. However, the task still remains within the understanding level because students are guided by explicit contextual clues and are not required to critically evaluate the text.

Another example appears in Item 3:

“The implied information that can be concluded from the text is...”

This item involves inference-making because students must derive information that is not directly stated in the text. Although inferential thinking reflects a deeper level of comprehension, the cognitive demand remains within LOTS because the reasoning process is still limited to interpreting textual meaning rather than analyzing relationships between ideas or evaluating arguments critically.

The most dominant cognitive level identified in the assessment is C3 (applying). Item 2 illustrates this category:

“The sentence that demonstrates the use of visual descriptive language in the text is...”

This question requires students to apply their understanding of descriptive language features in order to identify the correct sentence. The task reflects the implementing process because students must use conceptual knowledge within a specific context. Similarly, item 7 asks students to identify the correct use of an adverb of time in a sentence. These tasks demonstrate functional understanding of linguistic concepts; however, the questions remain procedural because the answers are guided by clearly identifiable criteria.

The predominance of C3 items suggests that teachers have attempted to move beyond simple memorization by encouraging students to apply concepts contextually. Nevertheless, these application tasks still do not fully engage students in analytical or evaluative reasoning. The assessment largely measures whether students can use previously learned concepts correctly rather than whether they can critique, synthesize, or construct new ideas.

This condition supports the findings of Anderson et al. (2023), who state that numerous classroom assessments still focus on procedural application rather than deeper cognitive engagement. Similarly, Mahendra (2023) argues that HOTS-oriented assessments remain difficult to implement because teachers tend to prioritize questions that are easier to construct and score objectively. As a result, students are rarely challenged to develop reflective and creative thinking abilities through classroom evaluation.

The interview results with the Indonesian language teacher further strengthen these findings. The teacher explained that test construction was mainly based on students' average academic abilities and the intention to ensure that the examination remained accessible. Consequently, teachers tend to avoid cognitively demanding tasks because they are perceived as too difficult for students. This finding suggests that the dominance of LOTS is influenced not only by teachers' understanding of assessment design but also by practical considerations related to students' performance and classroom conditions.

Overall, the dominance of LOTS in the examination instrument demonstrates that the assessment still emphasizes knowledge acquisition and procedural understanding. Although such cognitive levels remain necessary as foundational competencies, the lack of balance between LOTS and HOTS limits students' opportunities to develop critical, analytical, and creative thinking skills. Therefore, a more proportional integration of HOTS-oriented tasks is necessary to support the cognitive goals of the Merdeka Curriculum.

Representation of HOTS in Test Items

The analysis identified only three items categorized as HOTS, all of which were classified at the C4 (analyzing) level. No items were found at the evaluating (C5) or creating (C6) levels. This finding indicates that although the assessment includes several analytical tasks, the representation of higher-order thinking skills remains limited and incomplete.

One representative HOTS item is Item 19:

“The correct order of the sentences to form a coherent descriptive paragraph is...”

This item requires students to organize sentences into a coherent structure by identifying logical relationships between ideas. Unlike LOTS questions, students cannot answer the item merely by recalling factual information. Instead, they must analyze sentence coherence, identify relationships among textual elements, and determine the most logical sequence. The dominant cognitive process is organizing, which is one of the key processes within the analyzing dimension (C4).

Another HOTS item appears in Item 49:

“The correct sequence of the steps is...”

This question asks students to arrange procedural steps in the correct logical order. Students must analyze the relationships among the instructions and determine how each step contributes to the overall process. The task, therefore, requires structural reasoning and organizational thinking rather than simple memorization.

Similarly, item 43 asks students to determine the structural part of a procedural text. To answer the question correctly, students must differentiate textual functions and identify the role of the paragraph within the broader text structure. This process reflects analytical thinking because students must distinguish between several possible organizational components.

Although these items involve analytical reasoning, the HOTS representation remains limited because the cognitive demand does not progress beyond analysis. Students are not required to evaluate arguments, justify opinions, critique information, or create original responses. In other words, the assessment only partially reflects higher-order thinking processes.

This finding aligns with the argument of Surayanah and Karma (2023), who emphasize that HOTS-oriented assessment should involve evaluative and creative tasks rather than stopping at analytical processes. Similarly, Bisri et al. (2023) state that meaningful HOTS assessment requires students to engage in critical judgment, problem-solving, and idea generation. Without these components, assessments may still fail to stimulate deeper cognitive engagement.

The absence of C5 and C6 items is particularly significant because evaluative and creative thinking are central competencies in 21st-century learning. According to Karuru et al. (2023), assessments that lack opportunities for evaluation and creation tend to encourage passive learning patterns focused on reproduction of knowledge rather than innovation and critical inquiry. Therefore, the limited HOTS representation in this assessment suggests that students are not yet fully supported in developing the advanced cognitive skills required by contemporary educational frameworks.

Overall, the findings demonstrate that HOTS integration within the assessment instrument remains partial and underdeveloped. Although several items already encourage analytical reasoning, the absence of evaluative and creative tasks indicates that the assessment has not comprehensively measured higher-order thinking skills. Consequently, the examination instrument still falls short of the cognitive expectations emphasized in the Merdeka Curriculum.

The Alignment Between Test Items and Learning Objectives

The analysis of alignment between test items and learning objectives was conducted to examine the extent to which the assessment instrument reflects the competencies targeted in the learning objectives. This analysis refers to Webb’s alignment framework, which emphasizes consistency between instructional objectives, assessment content, and cognitive demands (Webb, 1997). In this study, alignment was examined through two main dimensions: content relevance and cognitive-level correspondence between the test items and the intended learning objectives.

Table 3. Alignment Between Test Items and Learning Objectives

Categories	Number of Items	Percentages
Aligned	32	64%
Not Aligned	18	36%
Total	50	100%

Table 3 above reveals that out of the 50 analyzed items, 32 items (64%) are categorized as aligned with the learning objectives, while 18 items (36%) are categorized as not aligned. These findings indicate that although most items adequately represent the intended competencies, a substantial proportion still fail to

correspond to the expected learning outcomes. This suggests that the quality of alignment within the assessment instrument has not yet been fully optimized. Similar findings are reported by Rahayu and Rosawati (2023), who found that numerous classroom assessments still demonstrate inconsistencies between intended competencies and measured cognitive demands.

In general, aligned items demonstrate a clear relationship between the instructional objectives and the cognitive processes measured in the assessment. These items are not only relevant to the instructional content but also reflect the cognitive demands expected in the curriculum. According to Biggs and Tang (2022), effective assessment should maintain constructive alignment between learning outcomes, teaching activities, and evaluation processes. When alignment is achieved, assessment functions not only as a measurement tool but also as a mechanism that supports meaningful learning.

One example of an aligned item is Item 1:

“The main purpose of the author in the text ‘The Instagrammable Pantan Terong’ is ...”

This item is categorized at the C2 (understanding) level because students are required to interpret the author’s communicative purpose based on the content of the descriptive text. The corresponding learning objective also emphasizes comprehension of descriptive texts. Therefore, the cognitive demand measured by the item is consistent with the intended instructional competency. This alignment indicates that the item functions as a valid measure of the targeted learning outcome. Anderson and Krathwohl explain that understanding-level tasks involve interpreting and constructing meaning from instructional messages, which is reflected in the cognitive demand of this item (Anderson & Krathwohl, 2015).

Alignment is also evident in Item 2:

“The sentence that demonstrates the use of visual descriptive language in the text is ...”

This question operates at the C3 (applying) level because students must apply their understanding of descriptive language features in order to identify the correct sentence. The related learning objective similarly requires students to recognize and use descriptive language contextually. Consequently, both the content and cognitive demand of the item correspond to the intended learning objective. This finding supports Zhou et al. (2023), who argue that assessments should encourage the contextual application of knowledge in order to support meaningful cognitive engagement and competency development.

Several other items also demonstrate alignment across cognitive levels ranging from C1 to C4. These questions generally reflect competencies related to understanding text content, identifying text structures, and applying language features contextually. This finding indicates that part of the assessment instrument has been designed with consideration of both instructional content and cognitive expectations. Nevertheless, the aligned items remain predominantly concentrated at lower cognitive levels, particularly C1–C3. As a result, although technical alignment has been achieved, the cognitive complexity of the assessment remains limited.

However, the study also identified a significant number of misaligned items. The primary issue concerns the mismatch between the cognitive level of the item and the level specified in the learning objectives. One example appears in Item 8:

“The description above involves which sense?”

This item is categorized at the C2 (understanding) level because students only need to identify sensory information in the text. However, the corresponding learning objective requires students to apply descriptive language features contextually at the C3 (applying) level. As a result, the item does not fully represent the expected competency because the cognitive demand remains below the intended level. According to Webb (1997), alignment requires equivalence not only in content but also in cognitive complexity. Therefore, items that measure lower cognitive processes than those specified in the objectives cannot be considered fully aligned.

A more substantial discrepancy appears in Item 22:

“Which of the following statements is a characteristic of folk poetry?”

This question only requires students to recall factual characteristics of folk poetry, placing it at the C1 (remembering) level. In contrast, the learning objective expects students to compare the types and structural

elements of folk poetry, which corresponds to the C4 (analyzing) level. The gap between these cognitive demands demonstrates that the item is unable to adequately measure the intended competency. Similar patterns were identified by Aydin and Birgili (2023), who found that assessment instruments often emphasize factual recall even when curricula require analytical reasoning.

These findings suggest that assessment practices still tend to prioritize factual mastery and basic comprehension rather than higher-order cognitive engagement. Teachers appear to focus more heavily on ensuring that students understand content at a surface level rather than encouraging analytical and reflective thinking. Consequently, although the test items may be content-relevant, they often fail to measure the complexity of thinking expected in the learning objectives. Mahendra (2023) explains that teachers frequently prefer lower-level questions because they are easier to construct and score objectively, even though such practices may reduce opportunities for developing higher-order thinking skills.

The findings also indicate that misalignment may influence the quality of learning itself. When students are primarily assessed through low-level cognitive tasks, they are more likely to focus on memorization and procedural understanding rather than developing analytical and critical thinking skills. In the long term, this condition may weaken the implementation of competency-based learning, which is a central principle of the Merdeka Curriculum. Tran et al. (2023) emphasize that assessments strongly influence students' learning behavior because learners generally adapt their study strategies to the cognitive demands presented in examinations.

Although the majority of items in this study were technically aligned with the learning objectives, the cognitive distribution remained heavily concentrated within LOTS categories. This means that alignment was often achieved only at lower cognitive levels, such as remembering, understanding, and applying. Consequently, the assessment instrument still provides limited opportunities for students to engage in evaluation, reasoning, and creative thinking processes. Surayanah and Karma (2023) argue that meaningful HOTS-oriented assessment should involve evaluative and creative dimensions in order to develop reflective and innovative thinking skills.

This finding reinforces Webb's argument that alignment should not merely concern similarities in content but must also involve equivalence in cognitive complexity (Webb, 1997). Therefore, assessment design should move beyond material relevance and pay greater attention to the level of thinking required in the learning objectives. In addition, Biggs and Tang (2022) state that assessment quality depends on the consistency between intended learning outcomes and the cognitive processes measured in evaluation activities. Overall, the results indicate that the alignment between test items and learning objectives can be categorized as moderately adequate but not yet optimal. The presence of 36% misaligned items demonstrates the need for improvement in assessment construction, particularly in ensuring consistency between cognitive levels and targeted competencies. Teachers need to systematically map cognitive processes using the Revised Bloom's Taxonomy so that assessment instruments not only measure content mastery but also promote students' critical, analytical, and creative thinking skills in accordance with the demands of 21st-century learning and the Merdeka Curriculum.

CONCLUSIONS AND RECOMMENDATION

Based on the research findings, the Final Semester Examination (UAS) items for Grade VII Indonesian Language at SMP Negeri 2 Dayun are still predominantly dominated by the Lower-Order Thinking Skills (LOTS) category. Of the 50 analyzed items, 47 items (94%) fall within the cognitive levels of C1 (remembering), C2 (understanding), and C3 (applying), with the highest concentration at the C3 level. In contrast, only 3 items (6%) are categorized as Higher-Order Thinking Skills (HOTS), all of which are limited to the C4 (analyzing) level, while no items are found at the C5 (evaluating) and C6 (creating) levels. In terms of alignment with learning objectives, 32 items (64%) are categorized as aligned, whereas 18 items (36%) are not aligned due to inconsistencies between the cognitive demands of the items and the competencies targeted in the learning objectives.

These findings indicate that the assessment instrument still primarily emphasizes factual mastery and basic cognitive processes such as remembering, understanding, and applying. As a result, students' analytical, evaluative, and creative thinking skills have not been optimally facilitated through the assessment process. The dominance of LOTS-oriented items demonstrates that the assessment practices implemented in this context have not yet fully reflected the demands of 21st-century learning and the competency-based orientation of the Merdeka Curriculum, which emphasizes critical, reflective, and creative thinking skills.

Theoretically, this study contributes to the field of educational assessment by strengthening the understanding that assessment quality should not only be viewed from the perspective of content relevance but also from the alignment of cognitive complexity between learning objectives and assessment instruments. This study expands previous research by integrating two analytical dimensions simultaneously, namely cognitive-level analysis based on the Revised Bloom's Taxonomy and alignment analysis based on Webb's alignment framework. Therefore, the findings provide a more comprehensive perspective on how cognitive distribution and alignment jointly influence the validity and effectiveness of classroom assessment practices. In addition, this study reinforces the argument that assessment instruments dominated by LOTS tend to limit the measurement of higher-order cognitive competencies expected in competency-based curricula.

Practically, this study provides essential implications for teachers, schools, and curriculum developers in improving the quality of classroom assessment. The findings reveal the need for more systematic assessment design that considers both cognitive-level distribution and consistency with learning objectives. Teachers are encouraged to develop assessment instruments that contain a more balanced proportion of LOTS and HOTS items so that students are not only trained to recall and understand information but are also encouraged to analyze, evaluate, and create. Furthermore, teachers need to ensure that each test item reflects the cognitive demands specified in the learning objectives to achieve more valid and representative assessment outcomes.

The research results also suggest the importance of professional development programs and training related to HOTS-based assessment construction. Through such training, teachers may improve their understanding of cognitive taxonomy, item-writing techniques, and alignment principles, thereby enabling them to design assessments that better support competency-based learning. In addition, schools and educational stakeholders may utilize the findings of this study as evaluative material in strengthening assessment policies aligned with the implementation of the Merdeka Curriculum.

Overall, the present study highlights that improving assessment quality requires not only attention to content coverage but also careful consideration of cognitive complexity and instructional alignment. Therefore, future assessment practices should be directed toward the development of evaluation instruments that are capable of comprehensively measuring students' critical, analytical, evaluative, and creative thinking skills in order to better support meaningful learning and the achievement of educational goals in the 21st century.

Limitation

The present study has limitations related to the number and type of instruments analyzed. The study only examines 50 items from the first semester final examination (UAS) of Grade VII Indonesian Language in a single school, which limits the generalizability of the findings. This means that the results regarding the dominance of LOTS and the alignment of test items with learning objectives may not necessarily represent the patterns of test items in other schools, different grade levels, or other subjects. If a larger number of items were analyzed or if samples were drawn from different semesters, the results of the cognitive level analysis and item alignment might also vary.

In addition, this study does not empirically evaluate the validity and reliability of the test items. The analysis is conducted solely based on content aspects and cognitive levels in accordance with Anderson and Krathwohl's theory and Webb's Alignment Model, without statistical testing or expert judgment to assess content and construct validity. As a result, the conclusions regarding the quality of the instrument remain descriptive and are not yet fully comprehensive. Consequently, although the study is able to illustrate the

dominance of LOTS and the theoretical alignment of items, it does not yet provide strong empirical evidence regarding the extent to which the test items validly and reliably measure students' thinking skills.

Credit authorship contribution statement

The first author is responsible for conducting research planning, data collection, and analysis, as well as writing the manuscript. The second author plays a role in supervision and conducts manuscript review.

Declaration of competing interest

The authors declare that they have no financial, professional, or personal conflicts of interest that could influence the results, interpretation, or reporting of the research presented in this paper.

Ethical Declaration

All participants provide informed consent before participating in the study. They were informed about the research objectives, procedures, and their right to withdraw at any time without any consequences or harm.

Acknowledgment

The author expresses gratitude to Almighty God for His grace and blessings, which have made this research possible. Sincere thanks are also extended to the author's beloved parents, Mr. Alfredo and Mrs. Gusnia Waty, for their prayers, support, and continuous motivation. The author greatly appreciates the guidance and direction provided by the academic supervisor, Mr. Alber, S.Pd., M.Pd., who has contributed significantly to the completion of this research. Furthermore, the author would like to thank the Principal of SMP Negeri 2 Dayun, all teachers especially the Indonesian language teacher and the administrative staff for granting permission and providing facilities for this study. Finally, the author extends appreciation to Dr. Muhammad Mukhlis, S.Pd., M.Pd., for his valuable input and suggestions in improving the quality of this research.

REFERENCES

- Abraham, I., Tjalla, A., & Indrajit, R. E. (2021). HOTS (Higher Order Thinking Skill) dalam pedagogik kritis. *JISIP (Jurnal Ilmu Sosial dan Pendidikan)*, 5(3).
- Aini, N., Rahman, F., & Putra, M. (2023). Teachers' challenges in developing HOTS-based assessment instruments. *Journal of Educational Research and Evaluation*, 7(2), 145–156.
- Aini, R., Karma, I. N., & Affandi, L. H. (2023). Kesulitan guru dalam menyusun soal evaluasi berbasis higher order thinking skills dalam pembelajaran Kurikulum 2013. *Jurnal Educatio FKIP UNMA*, 9(4), 2062–2069.
- Anderson, D. K., Schoenleber, M., & Korshavn, S. (2023). Higher-order clicker questions engage students and prepare them for higher-order thinking activities. *Journal of Microbiology & Biology Education*, 24(1). <https://doi.org/10.1128/jmbe.00151-22>
- Anderson, L. W., & Krathwohl, D. R. (2015). *Kerangka landasan untuk pembelajaran, pengajaran, dan asesmen: Revisi taksonomi pendidikan Bloom (Abridged ed.)*. Pustaka Pelajar.
- Armianti, A., Subhan, M., Nasution, M. L., Al Aziz, S., Rani, M. M., Rifandi, R., & Harisman, Y. (2020). Profesionalisme guru dalam membuat soal higher order thinking skills. *JNPM (Jurnal Nasional Pendidikan Matematika)*, 4(1), 75–84.
- Aydin, U., & Birgili, B. (2023). Assessing mathematical higher-order thinking skills: An analysis of Turkish university entrance examinations. *Educational Assessment*, 28(3), 190–209. <https://doi.org/10.1080/10627197.2023.2202311>
- Biggs, J., & Tang, C. (2022). *Teaching for quality learning at university (5th ed.)*. McGraw-Hill Education. <https://www.mheducation.com/highered/product/teaching-quality-learning-university-biggs-tang/M9780335242757.html>

- Bisri, H., Rahmawati, Y., & Ridwan, A. (2023). Inventory for measuring higher-order thinking learning environment. *The New Educational Review*, 73, 162–174. <https://doi.org/10.15804/tner.23.73.3.12>
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40.
- Darmiyati Zuchdi, E., & Afifah, W. (2021). Analisis konten, etnografi, grounded theory, dan hermeneutika dalam penelitian. Bumi Aksara.
- Karuru, P., Muta'allim, Suparjo, Setiawan, A. F., & Junaida, S. (2023). Improving students' higher order thinking skills through a question and answer method. *RETORIKA: Jurnal Ilmu Bahasa*, 9(3), 340–349. <https://doi.org/10.22225/jr.9.3.2023.340-349>
- Karo-Karo, R., Simanjuntak, T., & Gultom, E. (2025). Cognitive level distribution in Indonesian language assessment. *Jurnal Pendidikan Bahasa Indonesia*, 14(1), 55–68.
- Kementerian Pendidikan, Riset, dan Teknologi. (2022). Panduan pembelajaran dan asesmen. Kemendikbudristek.
- Khoiriyah, Z., & Shaleh, S. (2025). Solusi alternatif atas problematika dalam mengimplementasikan penilaian higher order thinking skills (HOTS). *Al-Madrasah: Jurnal Ilmiah Pendidikan Madrasah Ibtidaiyah*, 9(2), 656–667.
- Mahendra, E. (2023). Oral assessment in triggering student higher-order thinking skills (HOTS). *JISAE: Journal of Indonesian Student Assessment and Evaluation*, 9(1), 8–14. <https://doi.org/10.21009/jisae.v9i1.32928>
- Margareta, T., Hidayatullah, R. W., Qaulan, A., Hasfin, T., Wahyuningsih, L., Islam, U., Raden, N., & Said, M. (2024). Analysis of HOTS question items for Bahasa Indonesia final assessment for VIIth grade students at MTs Negeri 6 Boyolali, 8, 48–63.
- Moleong, L. J. (2017). Metodologi penelitian kualitatif. PT Remaja Rosdakarya.
- Nugroho, R. A. (2018). HOTS (Kemampuan berpikir tingkat tinggi): Konsep, pembelajaran, penilaian, dan soal-soal. PT Gramedia Widiasarana Indonesia.
- Nurrohman, S., & Mukhlis, M. (2023). Analysis of cognitive levels in Indonesian language examination items. *Jurnal Evaluasi Pendidikan*, 11(2), 88–101.
- Ramalia, A., & Syamsurizal, S. (2025). Tinjauan literatur implementasi asesmen dalam Kurikulum Merdeka. *Jurnal Pengabdian Masyarakat dan Riset Pendidikan*, 4(1), 2907–2912.
- Sugiyono. (2015). Metode penelitian pendidikan: Pendekatan kuantitatif, kualitatif, dan R&D. Alfabeta.
- Sukardi, H. M. (2008). Evaluasi pendidikan: Prinsip dan operasionalnya (F. Yustianti, Ed.). Bumi Aksara.
- Surayanah, S., & Karma, L. (2023). Developing higher-order thinking skills through local wisdom integration in science learning. *Jurnal Ilmiah Ilmu Pendidikan*, 6(5), 3469–3481. <https://doi.org/10.54371/jiip.v6i5.2054>
- Waruwu, M., Puat, S. N., Utami, P. R., Yanti, E., & Rusydiana, M. (2025). Metode penelitian kuantitatif: Konsep, jenis, tahapan, dan kelebihan. *Jurnal Ilmiah Profesi Pendidikan*, 10(1), 917–932.
- Webb, N. (1997). Criteria for alignment of expectations and assessments in mathematics and science education (Research Monograph No. 6).
- Yulistio, D., & Nugroho, A. (2024). Evaluasi pedoman (alat penilai) penyusunan karakteristik tes (soal) Bahasa Indonesia ranah kognitif berbasis HOTS. *Silampari Bisa: Jurnal Penelitian Pendidikan Bahasa Indonesia, Daerah, dan Asing*, 7(1), 31–49.
- Zhou, Y., Gan, L., Chen, J., Wijaya, T. T., & Li, Y. (2023). Development and validation of a higher-order thinking skills assessment scale for pre-service teachers. *Thinking Skills and Creativity*, 48, 101272. <https://doi.org/10.1016/j.tsc.2023.101272>
- Zulianto, S. (2021). Keterampilan berpikir tingkat tinggi dalam soal ujian nasional Bahasa Indonesia. *Jurnal Kependidikan: Penelitian Inovasi Pembelajaran*.